

KEKER
VAN NEST
& PETERS

LATHAM & WATKINS LLP

MORRISON FOERSTER

May 14, 2025

VIA ECF

Hon. Ona T. Wang
 Daniel Patrick Moynihan United States Courthouse
 500 Pearl Street
 New York, NY 10007

Re: Reconsideration of Order Directing OpenAI to Preserve Output Logs; *In Re: OpenAI, Inc., Copyright Infringement Litigation*, No. 1:25-md-3143, This Document Relates To: *The New York Times Company v. Microsoft Corporation, et al.*, No. 1:23-cv-11195

Dear Magistrate Judge Wang:

We write to respectfully request that the Court reconsider its order directing OpenAI to “preserve and segregate all output log data that would otherwise be deleted.” ECF 551 (“Order”).¹ The Order requires OpenAI to disregard legal, contractual, regulatory, and ethical commitments to hundreds of millions of people, businesses, educational institutions, and governments around the world—even though there is no reason to believe these drastic measures will advance this litigation. To be clear: OpenAI is taking the steps it can to comply with the Order despite the many practical and engineering challenges compliance entails. But we urge the Court to reconsider the Order, both to correct a manifest injustice and because the Order is based on material misrepresentations in Plaintiffs’ recent letter brief—to which OpenAI has not yet had the opportunity to respond. ECF 545 (“Letter Brief”).

First, the Order is contrary to law because there is no basis to believe that the data it covers is relevant or that preserving it will advance this litigation. OpenAI has retained *tens of billions* of user conversations. Plaintiffs have never even attempted to show that the retained data is not sufficiently representative to meet the needs of this case. There is no basis to conclude that the conversation data subject to the Order is relevant—as opposed to cumulative of the billions of conversations OpenAI has retained.

Second, the Letter Brief that prompted the Order contains two serious factual misstatements: (a) that OpenAI deleted user data in response to litigation events, and (b) that News Plaintiffs² only recently learned of OpenAI’s data retention practices. Both are false: (a) OpenAI has not deleted user conversation data for any reason other than to comply with its longstanding data practices, and (b) OpenAI informed News Plaintiffs’ counsel of those practices over 14 months ago. News Plaintiffs, in other words, elicited the Order by misrepresenting both OpenAI’s conduct and the

¹ ECF citations are to the docket in Case No. 1:23-cv-11195.

² “News Plaintiffs” refers to the New York Times and Daily News Plaintiffs; “Plaintiffs” includes the foregoing parties plus CIR.

factual record. That too warrants reconsideration.

* * *

OpenAI is taking the steps it can to comply with the Order. To be clear, compliance is not a switch-flip: OpenAI's data infrastructure is complex and retaining data that is otherwise slated for deletion requires significant engineering work, infrastructure changes, and compute resources. OpenAI has already commenced that work in light of the Order's requirements, though certain aspects of the company's services pose unique technical and infrastructural challenges. OpenAI stands ready to submit a sworn declaration with more detail as to the steps it has taken, as well as the aforementioned challenges.

OpenAI nevertheless requests that the Order be reconsidered because it is contrary to law, relies on News Plaintiffs' factual misstatements, and severely prejudices OpenAI and its users. These are more than sufficient grounds for reconsideration. *See generally In re Iraq Telecom Limited*, No. 18-mc-458-LGS-OTW, 2019 WL 5080007, at *1 (S.D.N.Y. 2019) ("major grounds justifying reconsideration" include "availability of new evidence[] or the need to correct a clear error or prevent manifest injustice" (citation omitted)).

First, preservation orders require, at minimum, some showing that the "documents in jeopardy are in fact relevant." *Treppel v. Biovail Corp.*, 233 F.R.D. 363, 370–71 (S.D.N.Y. 2006).³ And courts' "inherent authority to order those before them to preserve potentially relevant evidence . . . ordinarily should not be used in the absence of some substantial basis for believing that it would serve a useful purpose." *In re Application of Chevron Corp.*, 736 F. Supp. 2d 773, 787 (S.D.N.Y. 2010).

No such basis exists here. OpenAI now has to disregard its users' privacy interests based on Plaintiffs' unsupported speculation about what *might* exist, even though they have not even begun to analyze what *does* exist. Plaintiffs are well aware that wholesale preservation will not materially advance their claims because OpenAI is retaining *tens of billions* of conversations—*i.e.*, far more than any expert could review, and an amount Plaintiffs recognize must be analyzed through sampling. Plaintiffs have not presented any reason to believe that the data covered by the Order will provide some unique value that they cannot derive from the data OpenAI already retains. Under these circumstances, requiring OpenAI to disregard the privacy interests of its users is wholly disproportionate to the needs of the case and unwarranted.⁴

³ See also *Toussie v. Allstate Ins. Co.*, No. 15-cv-5235, 2018 WL 2766140, at *6 (E.D.N.Y. June 8, 2018), *aff'd*, 2018 WL 11451597 (E.D.N.Y. Aug. 15, 2018).

⁴ Plaintiffs also note that the number of conversation deletions was [REDACTED] in December 2023. But Plaintiffs focus on a time period when [REDACTED] (shortly before News Years' Day 2024). At almost every other point in time, the proportion of deleted ChatGPT conversations was [REDACTED]. Indeed, the chart Plaintiffs submitted indicates that, in December 2024, OpenAI was retaining [REDACTED] conversation data.

This is particularly concerning because the Order requires a massive disruption to commonplace data retention practices that have nothing to do with this litigation. News Plaintiffs have repeatedly characterized the data in question as data OpenAI “destro[ys].” But, in reality, the data OpenAI does not retain is (1) data that OpenAI’s *users* elect not to retain, including ChatGPT conversations users specifically choose to delete; and (2) data that OpenAI agrees not to retain under either its negotiated agreements or general business terms with its *customers*. OpenAI commits not to retain this data for a simple reason: because millions of individuals, businesses, and other organizations use OpenAI’s services in a way that implicates uniquely private information—including sensitive personal information, proprietary business data, and internal government documents.

The Order prohibits OpenAI from honoring this commitment by (*e.g.*) forbidding OpenAI from complying with a user’s attempt to delete a conversation about a family member’s health condition or immigration status. That will harm not only OpenAI, but also the hundreds of millions of users who rely on its services. That counsels in favor of reconsideration, particularly when the Order requires preservation of data that will not actually advance this litigation.⁵

Second, the Letter Brief on which the Order is based misstates the facts. It focuses on [REDACTED] and suggests that [REDACTED] reflect targeted deletions by OpenAI of user data in response to litigation events. This is unequivocally false: as OpenAI would have informed News Plaintiffs had they sought a meet-and-confer in advance of the Letter Brief’s filing, [REDACTED] were caused by [REDACTED] [REDACTED]. That much is evident from the chart Plaintiffs cite in their letter, which shows [REDACTED] [REDACTED] had anything to do with this litigation.

The Letter Brief also claims that News Plaintiffs learned of OpenAI’s data retention practices “for the first time” on November 15, 2024. Letter Brief at 2. Again, unequivocally false: OpenAI put counsel for the New York Times (including Rothwell Figg, who also represents the Daily News Plaintiffs) on notice of OpenAI’s retention policies as to both categories of data in a letter in February 2024. ECF 379-3. That was *almost 10 months* before News Plaintiffs raised any

⁵ See *Commentary on Proportionality in Electronic Discovery*, 11 Sedona Conf. J. 289, 291 (2010) (preservation burden “should be weighed against the potential value and uniqueness of the information”); *Treppel*, 233 F.R.D. at 372 (denying preservation order where, *inter alia*, no showing that order “is ‘not unduly burdensome’”).

purported concerns with these policies—which are also publicly available online—in December 2024. ECF 379-14. And CIR has *never* raised this concern until now.

If Plaintiffs believed that this data was essential (or even relevant) to their claims, they would have said so when OpenAI notified News Plaintiffs’ counsel of its privacy and retention policies over a year ago. That they did not is a tacit admission that this data falls far short of the requirements for a preservation order. *See infra*. It also suggests that their recent briefing on this issue has nothing to do with advancing this litigation and everything to do with distracting from its efficient adjudication. The Court should not indulge that self-serving litigation tactic at the expense of the privacy of hundreds of millions of OpenAI’s users. OpenAI requests that the Court vacate the Order or, in the alternative, stay the Order pending the May 27 hearing, and instruct the parties to continue their dialogue as to sampling the retained data for expert analysis.

KEKER, VAN NEST
& PETERS LLP

/s/ Edward A. Bayley

Edward A. Bayley*

LATHAM & WATKINS LLP

/s/ Elana Nightingale Dawson

Elana Nightingale Dawson

MORRISON &
FOERSTER LLP

/s/ Joseph C. Gratz

Joseph C. Gratz*

cc: All Counsel of Record (via ECF)

* All parties whose electronic signatures are included herein have consented to the filing of this document, as contemplated by Rule 8.5(b) of the Court’s ECF Rules and Instructions.